Predicting Imminent Aggression Onset in Minimally-Verbal Youth with Autism Spectrum Disorder Using Preceding Physiological Signals

Matthew S. Goodwin  
Northeastern University  
Boston, MA, USA  
m.goodwin@northeastern.edu

Ozan Özdenizci  
Northeastern University  
Boston, MA, USA  
oozdenizci@ece.neu.edu

Catalina Cumpanasoiu  
Northeastern University  
Boston, MA, USA  
cumpanasoiu.d@husky.neu.edu

Peng Tian  
Northeastern University  
Boston, MA, USA  
pengtian@ece.neu.edu

Yuan Guo  
Northeastern University  
Boston, MA, USA  
yuanee20@ece.neu.edu

Amy Stedman  
Maine Medical Center Research Institute  
Portland, ME, USA  
stedma@springharbor.org

Christine Peura  
Maine Medical Center Research Institute  
Portland, ME, USA  
peurac@mmc.org

Carla Mazefsky  
University of Pittsburgh  
Pittsburgh, PA, USA  
mazefskyca@upmc.edu

Matthew Siegel  
Maine Medical Center Research Institute  
Portland, ME, USA  
siegem@springharbor.org

Deniz Erdoğmuş  
Northeastern University  
Boston, MA, USA  
erdogmus@ece.neu.edu

Stratis Ioannidis  
Northeastern University  
Boston, MA, USA  
ioannidis@ece.neu.edu

ABSTRACT
We test the hypothesis that changes in preceding physiological arousal can be used to predict imminent aggression proximally before it occurs in youth with autism spectrum disorder (ASD) who are minimally verbal (MV-ASD). We evaluate this hypothesis through statistical analyses performed on physiological biosensor data wirelessly recorded from 20 MV-ASD youth over 69 independent naturalistic observations in a hospital inpatient unit. Using ridge-regularized logistic regression, results demonstrate that, on average, our models are able to predict the onset of aggression 1 minute before it occurs using 3 minutes of prior data with a 0.71 AUC for global, and a 0.84 AUC for person-dependent models.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

PervasiveHealth ’18, May 21–24, 2018, New York, NY, USA  
© 2018 Association for Computing Machinery.  
ACM ISBN 978-1-4503-6450-8/18/05 . . . $15.00  
https://doi.org/10.1145/3240925.3240980

1 INTRODUCTION
Unpredictable and potentially dangerous aggressive behavior towards others by youth with autism spectrum disorder (ASD) isolates them from foundational educational, social,
Table 1: Naturalistic data collection descriptive statistics of participants.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Sessions</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>69</td>
<td>3.45</td>
</tr>
<tr>
<td>Total Obs. Duration*</td>
<td>9.33</td>
<td>3.97</td>
<td>3.77</td>
<td>3.02</td>
<td>1.25</td>
<td>0.57</td>
<td>7.02</td>
<td>8.47</td>
<td>2.72</td>
<td>1.43</td>
<td>0.22</td>
<td>1.38</td>
<td>1.6</td>
<td>8.47</td>
<td>0.52</td>
<td>1.02</td>
<td>20.48</td>
<td>0.78</td>
<td>5.1</td>
<td>5.87</td>
</tr>
<tr>
<td>Number of Aggression Episodes</td>
<td>72</td>
<td>7</td>
<td>13</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>35</td>
<td>30</td>
<td>50</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>9</td>
<td>39</td>
<td>1</td>
<td>8</td>
<td>130</td>
<td>2</td>
<td>76</td>
<td>47</td>
</tr>
<tr>
<td>Mean Agg. Duration†</td>
<td>9</td>
<td>102</td>
<td>77</td>
<td>11</td>
<td>19</td>
<td>9</td>
<td>19</td>
<td>18</td>
<td>50</td>
<td>3</td>
<td>1</td>
<td>15</td>
<td>19</td>
<td>6</td>
<td>51</td>
<td>7</td>
<td>103</td>
<td>4</td>
<td>7</td>
<td>22</td>
</tr>
</tbody>
</table>

* Total observation durations are presented in hours. † Mean aggression durations are presented in seconds.

and familial activities, thereby markedly exacerbating morbidity and costs associated with the condition. As many as 2/3 of youth with ASD display aggression [21], which is one of the primary reasons they use behavioral healthcare services [4]. Aggression presents imminent safety risks for the individual and others in the environment. It frequently co-occurs with agitation, meltdowns, and other problem behaviors that are difficult to manage. Families report that aggression increases their stress, isolation, and financial burden, and decreases available support options [9, 18]. Cross-sectional [11] and longitudinal studies [14, 42] in ASD suggest that even when broadly-defined problem behaviors decline, they remain heightened in comparison to typically developing (TD) and intellectually disabled (ID) populations; and some ASD subgroups engage in persistent or increasing problem behaviors in adulthood.

Aggression to others is particularly impairing and treatment refractory in the 30 – 40% of youth with ASD who are minimally verbal (MV-ASD). Their inability to self-report distress can lead to behaviors that seem to occur without warning, sometimes long after any observable trigger. This unpredictability makes aggression in MV-ASD particularly dangerous and is a barrier to accessing the community, medical providers, and educational placements as caregivers become afraid to put their child with ASD into potentially stressful environments that might lead to aggression to others. This predicament can demoralize caregivers, accelerate negative patient trajectories, decrease quality of life, and collectively increase health care costs.

A new approach to understanding and reducing aggression to others in ASD is needed, particularly for those with MV-ASD. Aggression is frequently treated with medication, which can have significant side effects and inconsistent success [1, 40, 41]. Another common evidence-based approach is to attempt to determine the function of aggressive behavior (escaping from demands, gaining attention, etc.) utilizing applied behavioral analysis (ABA) [5]. While ABA interventions that target these functions in ASD are common [19], their effectiveness is often reduced due to insufficient time to attempt de-escalation strategies before aggression occurs. Moreover, it has been shown that 30% of functional behavior assessment studies are inconclusive, and many with ASD remain aggressive even when the identified trigger is withdrawn or the apparent function is addressed [10, 20, 37].

Aggression to others may represent a maladaptive attempt to express or modulate physiological arousal due to distress [7]. In TD youth, greater ability to regulate physiological arousal is associated with fewer behavior problems [6, 31]. Studies of disorders characterized by emotional and behavioral dysregulation, such as bipolar disorder and antisocial behavior, report a strong association between physiological disturbance and symptomatology [12, 25, 29, 30, 32, 35, 38]. Prior research also demonstrates an association between physiological arousal and problem behavior in ASD [8, 13], wherein an individual may engage in a problem behavior in an attempt to communicate or alleviate distress, and decrease or increase arousal to achieve autonomic equilibrium [15–17, 33, 34]. If these behaviors are punished or their function is not satisfied, physiological arousal can increase, exacerbating and perpetuating an escalating loop of distress, arousal, and aggression. Thus, we hypothesize that changes in physiological arousal precede aggressive behavior. Our objective is to reduce the impact of aggression in MV-ASD by validating physiological biomarkers that precede the proximal onset of aggression to others. If successful, this effort could transform the approach to aggression in ASD by enabling providers in the future to receive real-time alerts predicting when aggression is imminent, creating new opportunities for preventing or mitigating its emergence, occurrence, and impact.
Predicting Aggression in Minimally-Verbal Youth with ASD

2 NATURALISTIC OBSERVATIONS
Participants
Twenty ADOS-2 [26] confirmed, behaviorally unstable, MV-ASD youth in the Developmental Disorders Unit at Spring Harbor Hospital in Portland, Maine participated in this study. The sample on average was 10.8 years old (SD = 3.10, range = 6 – 17), 75% male (n = 15), 95% Caucasian (n = 19), and 90% non-Hispanic (n = 18). Mean nonverbal IQ (NVIQ) of the sample assessed at hospital admission was 66.1 (SD = 18.99, range = 31 – 98), with 65% of patients (n = 13) meeting criteria for intellectual disability (NVIQ ≤ 70). Severity of behavior in the sample was also assessed at intake using the caregiver-rated Aberrant Behavior Checklist (ABC [3]) Irritability subscale, and resulted in a mean of 28.11 (SD = 8.55, range = 16 – 42).

Data Acquisition
Data recorded in this Institutional Review Board approved study included 69 unstructured observational sessions in the specialized ASD psychiatric inpatient unit. Patients wore an E4 biosensor on their wrists (Empatica Inc., United States) while inpatient research staff concurrently coded naturalistic observations of operationally defined aggression to others (e.g., hitting, kicking, biting, scratching, grabbing, pulling). Inter-rater reliability for observed aggression onset and offset between two research staff from the inpatient site yielded 0.90 percent agreement and a corresponding Cohen’s Kappa of 0.79, wherein a maximum tolerance of 2sec onset or offset difference between raters was considered an agreement (see aggression duration mean and standard deviation in Table 1, i.e., average duration of observed aggression in the sample was 28sec, much longer than a 2sec difference between raters) in 20% of data randomly selected from our corpus. Research staff conducted these data collection observations with minimal interference to participants’ regularly scheduled daily routines over the course of their inpatient stay. During school hours this consisted of academic lessons, free time, lunch in the cafeteria, and group activities (gym, music, etc.). On the inpatient unit, this consisted of various group activities, down time, on the playground, and while eating dinner. Since our data collection sessions were observational and naturalistic, researchers did not interact with participants before, during, or after aggression episodes, they only provided clinical staff with assistance donning and doffing the E4.

All 20 MV-ASD youth tolerated the E4 sensor after desensitization and usable data was obtained in all cases. Sixty-nine independent naturalistic observational sessions were collected (Table 1), [m(sd) = 3.45(2.84) observational sessions per patient], totaling 87hrs [m(sd) = 4.35hrs(4.8hrs) per patient]. Out of 548 total aggressions observed with concurrent E4 data, mean and standard deviation of aggression frequency and duration was 27(34) episodes in a four-hour period of 28sec(32sec) average length, respectively.

The following autonomic nervous system indices were recorded by the wrist-worn E4 biosensor to capture measures of physiological arousal: (1) heart rate and heart rate variability, which is a measure of the variation in beat-to-beat interval, both derived from blood volume pulse (BVP) and inter-beat interval (IBI) via photoplethysmography [2] at 64 Hz; and (2) electrodermal activity (EDA) sampled at 4 Hz, which reflects autonomic innervation of sweat glands and provides a sensitive measure of alterations in physiological arousal. To quantify changes in physical activity, the E4 records movement acceleration (ACC) using an embedded 3-axis accelerometer at 32 Hz sampling rate.

3 METHODS
Time-Series Feature Extraction
Naturalistic observations yielded labeled time-series data through 6 signal sources (i.e., BVP, IBI, EDA, ACCx, ACCy, ACCz). Statistical analyses on these physiological and physical activity signals were performed through extracted time-series features offline. In bins of 15 seconds, the following features were calculated: first, last, maximum, minimum, mean and median value, amount of unique values, and the sum, standard deviation and variance of values falling in a bin. In order to exploit temporal information of aggression episodes, we extracted two more features using time-synchronized binary aggression labels offline; time since past aggression (TPA), which indicates the amount of time elapsed since the last observation of an aggression episode; and a binary aggression observation flag (AOF) feature, indicating whether an aggression episode has so far occurred within that recording session. The standard deviation of each calculated feature across predictor time-series bins was included in all prediction models.

Logistic Regression Classifier Model
The predictability of aggressive behavior onset using extracted features was investigated through a ridge-regularized logistic regression for binary decision making in time. In particular, at every time point t, using features extracted in a previous time range [t − τp, t], a classifier was used to predict a binary dependent variable l, estimating whether aggression will be observed or not in an upcoming time range (t, t + τf].

We adopted a 5-fold cross-validation protocol to generate training and testing data splits, and repeated five times to produce confidence intervals. At each fold, the classifier was constructed with the training split through maximum
likelihood estimation for optimal ridge-regularized regression weights \( \beta = [\beta_0, \beta_1, \ldots, \beta_d]^T \), where \( d \) is the number of features. To predict, the classifier generates probabilities for two classes \( l = +1 \) (i.e., aggression) and \( l = -1 \) (i.e., non-aggression) in the form:

\[
P(l = +1|\mathbf{x}; \beta) = \frac{1}{1 + e^{-\beta^T\mathbf{x}}},
\]

where \( \mathbf{x} = [1, x_1, \ldots, x_d]^T \) corresponds to the concatenated feature vector from \([t - \tau_p, t]\). Receiver Operator Characteristic (ROC) curves and corresponding Area Under the Curve (AUC) values were calculated depending on the decision thresholds over these probabilities.

Among all extracted features, five different feature subsets were used as predictor variables (\( \mathbf{x} \)) in our analyses: (1) only temporal information (\( TPA, AOF \)); (2) only physical activity (\( ACC \)); (3) only physiological activity (\( BVP, IBI, EDA \)); (4) physical and physiological activity features combined (\( BVP, IBI, EDA, ACC \)); and (5) all extracted features combined (\( BVP, IBI, EDA, ACC, TPA, AOF \)).

4 RESULTS

Classification analyses for particular values of \( \tau_p \) and \( \tau_f \) were performed through both global and person-dependent models. In the former, time-series data across all sessions and all participants were concatenated, whereas in the latter data were pooled across sessions within each person. Data was processed for decision making every 15 seconds, which resulted in 20,863 samples in global prediction models.

Global Prediction Models

Global prediction models of aggression in the upcoming one minute using all extracted features from the past \( \tau_p = 180 \) seconds, which was the highest value the shortest individual session data permitted, resulted in the blue solid ROC curve presented in Figure 1 with a corresponding AUC value of 0.71. Figure 2(a) depicts increases in AUC when E4 biosensor data is included, compared to using temporal information on aggression episodes alone. Similar analyses for \( \tau_p = 60 \) seconds are presented in Figure 2(b).

Regarding the relationship between past time range (\( \tau_p \)) and future time range used to make aggression onset predictions (\( \tau_f \)), Figure 2(c) depicts stationary performance in global prediction models using all features from various past \( \tau_p \) durations. However, when using physical and physiological biosensor data from the past (c.f. bold dashed red curve), we observe relative AUC increases compared to temporal data of past aggression only (c.f. light dotted dark blue curve). Similar performance dynamics were observed when we increased feature space dimensionality with \( \tau_p \), as illustrated in Figures 2(a) and 2(b).

Person-Dependent Prediction Models

We repeat the above analysis with parameters \( \tau_p = 180 \) seconds and \( \tau_f = 60 \) seconds in person-dependent models; Table 2 contains the corresponding mean AUCs. Across participants, we observe an average 0.15 increase in AUC values up to 0.84 using E4 biosensor data compared to using temporal aggression information only. Figure 1 shows ROC curves corresponding to person-dependent model prediction performance using all features compared to the global prediction model with the same parameters. We observe a higher mean AUC and a much favored behavior of steep increases of sensitivity for low false positive rates in person-dependent models compared to global prediction models.

5 DISCUSSION

We demonstrate, for the first time, that naturalistically observed aggressive behavior in MV-ASD youth in a hospital inpatient setting can be predicted with high accuracy using proximal physiological and physical activity biosensor data and temporal information on recently observed aggressive episodes. While a growing number of researchers have used physiological data to discriminate affective states, anxiety, and challenging behaviors generally in autism [22–24, 27, 39], all prior work we are aware of in this area relies on artificial experimental settings and tasks, does not evaluate a sizeable
Figure 2: Mean AUC values of the global models varying as a function of: (a) $\tau_f$ using features from the past three minutes; (b) $\tau_f$ using features from the past minute; and (c) $\tau_p$ at $\tau_f = 60$ seconds. Each color represents one of the five feature subsets for all figures. Error bars represent minimum and maximum AUC values across cross-validation repetitions.

Table 2: Mean AUC values across cross-validation repetitions in person-dependent prediction models with $\tau_p = 180$ seconds and $\tau_f = 60$ seconds (i.e., predicting aggression onset for the next minute, using accumulated data from the past three minutes).

<table>
<thead>
<tr>
<th>Features</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
<th>P10</th>
<th>P11</th>
<th>P12</th>
<th>P13</th>
<th>P14</th>
<th>P15</th>
<th>P16</th>
<th>P17</th>
<th>P18</th>
<th>P19</th>
<th>P20</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal</td>
<td>0.67</td>
<td>0.57</td>
<td>0.58</td>
<td>0.52</td>
<td>0.83</td>
<td>0.58</td>
<td>0.62</td>
<td>0.58</td>
<td>0.83</td>
<td>0.82</td>
<td>0.61</td>
<td>0.64</td>
<td>0.59</td>
<td>0.92</td>
<td>0.83</td>
<td>0.65</td>
<td>0.99</td>
<td>0.53</td>
<td>0.61</td>
<td>0.69</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>0.66</td>
<td>0.84</td>
<td>0.80</td>
<td>0.49</td>
<td>0.88</td>
<td>0.73</td>
<td>0.65</td>
<td>0.79</td>
<td>0.98</td>
<td>0.51</td>
<td>0.85</td>
<td>0.79</td>
<td>0.64</td>
<td>0.81</td>
<td>0.73</td>
<td>0.72</td>
<td>0.99</td>
<td>0.60</td>
<td>0.75</td>
<td>0.75</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Physiological</td>
<td>0.64</td>
<td>0.85</td>
<td>0.79</td>
<td>0.68</td>
<td>0.99</td>
<td>0.68</td>
<td>0.66</td>
<td>0.70</td>
<td>0.88</td>
<td>0.97</td>
<td>0.78</td>
<td>0.82</td>
<td>0.79</td>
<td>0.70</td>
<td>0.97</td>
<td>0.70</td>
<td>0.75</td>
<td>0.99</td>
<td>0.69</td>
<td>0.74</td>
<td>0.79</td>
<td>0.12</td>
</tr>
<tr>
<td>Physical and Physiological</td>
<td>0.71</td>
<td>0.86</td>
<td>0.81</td>
<td>0.67</td>
<td>0.98</td>
<td>0.75</td>
<td>0.68</td>
<td>0.76</td>
<td>0.88</td>
<td>0.99</td>
<td>0.83</td>
<td>0.87</td>
<td>0.85</td>
<td>0.71</td>
<td>0.99</td>
<td>0.72</td>
<td>0.78</td>
<td>0.99</td>
<td>0.70</td>
<td>0.78</td>
<td>0.82</td>
<td>0.11</td>
</tr>
<tr>
<td>All Features Combined</td>
<td>0.74</td>
<td>0.87</td>
<td>0.81</td>
<td>0.69</td>
<td>0.98</td>
<td>0.77</td>
<td>0.69</td>
<td>0.77</td>
<td>0.92</td>
<td>0.99</td>
<td>0.86</td>
<td>0.85</td>
<td>0.91</td>
<td>0.73</td>
<td>0.99</td>
<td>0.89</td>
<td>0.80</td>
<td>0.99</td>
<td>0.71</td>
<td>0.78</td>
<td>0.84</td>
<td>0.10</td>
</tr>
</tbody>
</table>

sample of ends users in naturalistic settings, and is correlational, not temporally predictive. Our results demonstrate proof-of-concept, feasibility, and incipient validity predicting imminent aggression in this population and setting using both global and person-dependent models. As communicated to us by inpatient clinical staff, the potential benefit of avoiding or reducing a dangerous aggressive event is likely to outweigh the potential harm associated with a false positive in clinical practice. Predicting aggression onset during naturalistic observation in the upcoming 1 minute with at least 80% sensitivity is of high clinical value and could create new opportunities for preventing or mitigating aggression emergence, occurrence, and impact in MV-ASD.

We focused on the MV-ASD population because they are the most treatment-refractory, which compounds their underserved status in research [36]. We also focused on physical aggression to others because it is the most interfering and relevant (vs. verbal aggression) for MV-ASD youth. Finally, we sought to define a more generalized, objective approach to aggression grounded in underlying physiological mechanisms rather than trying to differentiate proactive or reactive aggression as is often done in the literature addressing the typically developing population. The advantages of biologically-based tools for predicting and preventing aggression are numerous, and the potential applications of...
these tools for improving the daily lives of individuals with MV-ASD are immense.

Our findings lay the groundwork for future work that defines and enables new opportunities for intervention before distress escalates to aggression. For example, with relatively modest ongoing technical development, biosensor data could be transmitted to a mobile phone via Bluetooth using a real-time data streaming application. The biosensor data and time-synchronized coding of aggression on the mobile application could be sent to a secure cloud server via wifi to store data, run classifiers, and push real-time aggression risk predictions via wifi or Bluetooth to a mobile phone, displaying a risk alert on the mobile application when indicated. Direct care staff could monitor these alerts and initiate de-escalation or emotion regulation interventions before aggression occurs. The impact of these alerts on the effectiveness of de-escalation and/or emotion regulation interventions could then be tested through a randomized controlled trial to determine if being alerted to imminent aggression (including the impact of false positive and false negative prediction rates) reduces the frequency, intensity, and duration of such episodes.

The impact of our work could be further extended by assessing whether degree of emotion regulation impairment is associated with specific arousal patterns, and if so, use that information to optimize subject selection for future intervention trials. We could also apply our methods to other problem behaviors in MV-ASD (self-injury, property destruction, elopement, etc.), including extending it to verbal youth, youth with other developmental disorders, and other populations who exhibit frequent aggression. In addition, this work could be extended through future treatment trials of an integrated mobile application and biosensor system with caregivers and in community settings, placing a validated biologically-based system for prediction of imminent aggression in their hands. Finally, with user interface modifications, this technology has the potential to be used by individuals with MV-ASD as a tool for self-monitoring and reminder to attempt self-regulation strategies.

Regarding future work currently underway, we will test whether hybrid models improve prediction performance, wherein person-dependent models are iteratively modified to include the most significant physiological biomarker features from global models. We will also explore ways to create more robust global prediction models, wherein we will test whether aggression onsets can be modeled as a non-homogeneous Poisson process. In that regard, we will evaluate whether regressing hazard rates from past observations can be performed through maximum likelihood estimation, comprising an iterative solution of several weighted linear regressions [28], leading to improved predictive performance across upcoming time ranges.

In sum, we seek through this work to combine novel methodologies, cutting-edge technology, unique settings, and focused multidisciplinary expertise to rigorously engage a historically intractable problem for an ASD population that suffers from the greatest morbidity, and who arguably are the most in need of innovative approaches.

ACKNOWLEDGMENTS

This work was supported by grants from the Simons Foundation (SFARI 296318), the Nancy Lurie Marks Family Foundation, the National Institute on Deafness and Other Communication Disorders (P50 DC013027), and NSF (SCH-1622536, IIS-1118061).

REFERENCES


Predicting Aggression in Minimally-Verbal Youth with ASD


